Chapter 12

Image Segmentation and Parameterization for Automatic Diagnostics of Whole-Body Scintigrams: Basic Concepts

Luka Šajn and Igor Kononenko

Contents

Bone scintigraphy or whole-body bone scan is one of the most common diagnostic procedures in nuclear medicine. Since expert physicians evaluate images manually, some automated procedure for pathology detection is desired. Scintigraphy segmentation into the main skeletal regions is briefly presented. The algorithm is simultaneously applied on anterior and posterior whole-body bone scintigrams. The expert's knowledge is represented as a set of parameterized rules, used to support image processing algorithms. The segmented bone regions are parameterized with algorithms for classifying patterns so the pathologies can be classified with machine learning algorithms. This approach enables automatic scintigraphy evaluation of pathological changes; thus, in addition to detection of pointlike high-uptake lesions, other types can be discovered. We extend the parameterization of the bone regions with multiresolutional approach and present an algorithm for image parameterization using the association rules.

Our study includes 467 consecutive, nonselected scintigrams. Automatic analysis of whole-body bone scans using our segmentation algorithm gives more accurate and reliable results than previous studies. Preliminary experiments show that our expert system based on machine learning closely mimics the results of expert physicians.

12.1 Introduction

The whole-body scan or the whole-body bone scintigraphy is a well-known clinical routine investigation and one of the most frequent diagnostic procedures in nuclear medicine [1]. Indications for bone scintigraphy include benign and malignant diseases, infections, degenerative changes, and other clinical entities [2]. Bone scintigraphy has high sensitivity, and the changes of the bone metabolism are seen earlier than the changes in bone structure detected on skeletal radiograms [1].

The investigator's role is to evaluate the image, which is of technically poor resolution due to the physical limitations of gamma camera. There are approximately 158 bones visible on anterior and posterior whole-body scans [3]. Poor image resolution and the number of bones to recognize make the evaluation of images difficult. Some research on automating the process of counting the bone lesions has been done, but only a few studies attempted to automatically segment individual bones prior to the computerized evaluation of bone scans [4–6].

12.1.1 Related Work

First attempts to automate scintigraphy in diagnostics for thyroid structure and function were made in 1973 [7]. Most of the research on automatic localization of bones was performed at the former Institute of Medical Information Science at the University of Hildesheim in Germany from 1994 to 1996. The main contribution was made by Bernauer [4] and Berning [5], who developed semantic representation of the skeleton and evaluation of the images. Benneke [6] realized their ideas in 1996.

Yin and Chiu [8] tried to find lesions using a fuzzy system. Their preprocessing of scintigrams includes rough segmentation of six fixed-size regions, regardless of individual image properties. Those parts are rigid and not specific enough to localize a specific bone. Their approach for locating abnormalities in bone scintigraphy is limited to pointlike lesions with high uptake.

When dealing with lesion detection, other authors, such as Noguchi [3], have been using intensity thresholding and manual lesion counting or manual bone region of interest (ROI) labeling. Those procedures are sufficient only for more evident pathologies, whereas new emerging pathological regions could be overlooked.

12.1.2 Our Aim and Approach

The aim of our study is to develop a robust procedure for diagnosing whole-body bone scans. Some possible methods for individual bone extraction are also presented. Segmented scans allow further development of automated procedures for recognition of pathological condition in specific bone regions. The experience with segmentation and pathology classification is presented. The steps of segmentation and diagnosing are shown in Figure 12.2.

12.1.2.1 Segmentation

When a scintigraphy is observed by an expert physician, each bone region is diagnosed according to several possible pathologies (lesions, malignom, metastasis, degenerative changes, inflammation, other pathologies, or no pathologies). The process of detecting the lesions can be aided by some advanced machine learning classifier [9] that produces independent diagnoses. The implementation of such a system was also the goal of the study. It can be used as a tool to remind a physician of some possibly overlooked spots or even to give some additional insight into the problem. It also enables further studies of individual bone regions with other algorithms (e.g., pattern classification or custom tailored algorithms).

In order to achieve a robust segmentation algorithm, we defined the most characteristic bone regions and the most identifiable points in those regions (reference points, Figure 12.1). Those points are chosen so that they can be uniformly identified over all distinct images in as many cases as possible, which is necessary since the images and skeletons vary considerably.

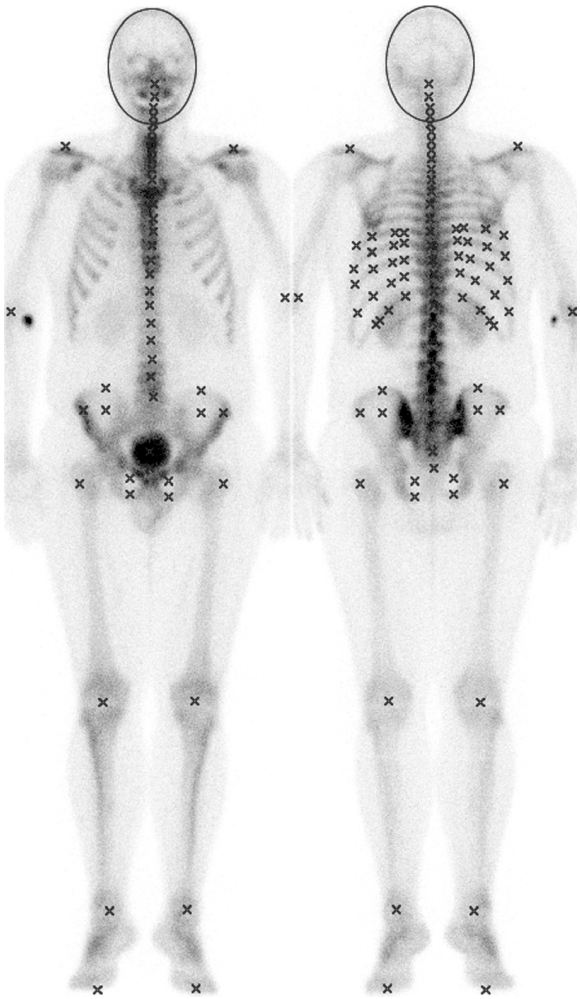


FIGURE 12.1: Defined characteristic reference points on anterior (left) and posterior (right) whole-body bone scintigrams.

There are some algorithms for detecting image features (e.g., scale-invariant feature transform, or SIFT [10]). Since the scintigrams are so variable, we cannot rely solely on features detected in images alone but must use some existing background knowledge. In our case, this background knowledge is the human anatomy.

Because scintigrams are represented by relatively small images, many algorithms (e.g., PCA [11], image correlation) are not appropriate or directly applicable. The idea is to use simple and easy-to-control algorithms. In our study, we used several image processing algorithms, such as binarization,

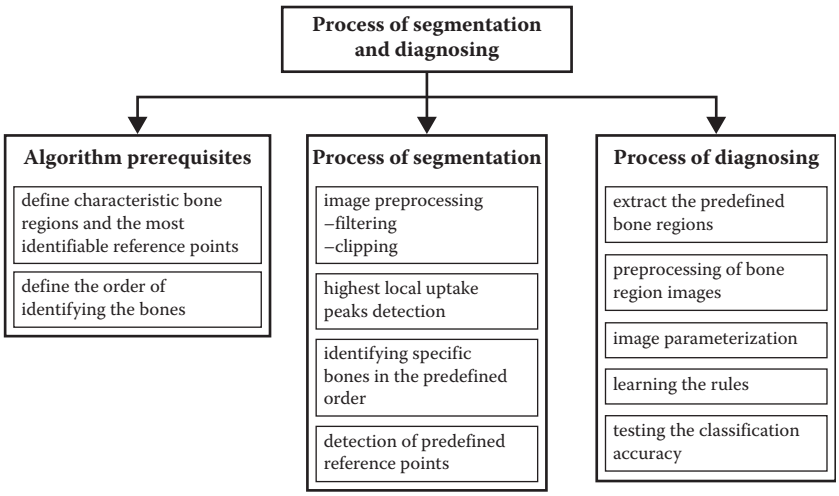


FIGURE 12.2: Algorithm steps.

dilatation, skeletonization, Hough's transform, Gaussian filtering [12], beam search, circle fitting, and ellipse fitting with least square method (LSM), in combination with background knowledge of anatomy and scintigraphic patterns.

12.1.2.2 Image Parameterization

In order for images to be automatically analyzed, they must be described in terms of a set of (numerical) parameters. There exist many different approaches to characterize textures of images. Most texture features are based on structural, statistical, or spectral properties of the image. For the purpose of diagnosis from medical images, it seems that structural description is most important. For that purpose, we use algorithm ArTeX for textural attributes that are based on association rules. The association rules algorithms can be used for describing textures if an appropriate texture representation formalism is used. This representation has several good properties, such as invariance to global lightness and invariance to rotation. Association rules capture structural and statistical information and are very convenient to identify the structures that occur most frequently and have the most discriminative power.

Another issue is the selection of the appropriate resolution for extracting most informative textural features. Image parameterization algorithms use descriptors of some local relations between image pixels where this search perimeter is bounded to a certain size. This means that they can give different results at different resolutions. The resolution used for extracting parameters is important and depends on the properties of the observed domain.

12.1.2.3 Image Classification

The ultimate goal of the medical image analysis is a decision about the diagnosis. When images are described with (hopefully) informative numerical attributes, we can use various machine learning algorithms [13] for generating a classification system. For that purpose, we can use decision trees, the naive Bayesian (NB) classifier, neural networks, k-nearest neighbors, support vector machines, and others. Our experience in a similar problem of diagnosing the ischemic heart disease from scintigraphic images of the heart muscle [14] is that the NB classifier gives best results. It turned out that also for diagnosing (parts of) bone scintigrams, as described in this chapter, the NB classifier performs the best. Therefore, we describe results obtained by the NB. It is also well known that in other problems of medical diagnosis, the NB usually outperforms other classifiers [15].

This chapter is organized as follows. First, the medical domain and the algorithm for segmenting whole-body bone scans are presented. Then, the algorithm ArTeX for pattern parameterization using association rules with an extended parameter set is described. The algorithm ARes for selecting the appropriate resolutions for better pattern description follows. Finally, the segmentation and pathology diagnosing results are presented.

12.2 Materials and Methods

12.2.1 Patients and Images

Retrospective review of 467 consecutive, nonselected scintigraphic images from 461 different patients investigated at the Nuclear Medicine Department, University Medical Centre in Ljubljana, Slovenia, from October 2003 to June 2004, was performed. Images were not preselected, so the study included a representative distribution of patients coming to examinations over the 9-month period. Images contained some artifacts and nonosseous uptake, such as urine contamination and medical accessories (i.e., urinary catheters) [16]. In addition, site of radiopharmaceutical injection is frequently visible (obstructs the image).

Twenty-one percent of the images were diagnosed as normal with no artifacts, meaning that no pathologies and no artifacts were present. Sixty percent of the images were diagnosed with slight pathology regardless of artifacts, 18% with strong pathology, and 1% of the images were classified as super-scans. Super-scan is obtained when the patient has very strong pathologies, which absorb most of the radiopharmaceutical so the other bone regions are vaguely visible.

In 18% of the images, partial scans (missing a part of the head or upper/lower extremities in the picture) were acquired, which complicates

the segmentation process. There were also children/adolescents with growth zones (5% of the images), manifested as increased osteoblastic activity in well-delineated areas with very high tracer uptake.

12.2.2 Bone Scintigraphy

All patients were scanned with gamma camera model Siemens Multi-SPECT with two detectors, equipped with LEHR (low-energy high-resolution) collimators. Scan speed was 8 cm per minute with no pixel zooming. Technetium-labeled phosphonate (99 m-Tc-DPD, Techneos®) was used. Bone scintigraphy was obtained 2 to 3 hours after intravenous injection of 750 MBq of radiopharmaceutical. The whole-body field was used to record anterior and posterior views digitally with resolution of 1024×256 pixels (approx. $205\,\text{cm} \times 61\,\text{cm}$). Images represent counts of detected gamma decays in each spatial unit with 16-bit grayscale depth.

12.2.3 Detection of Reference Points

Bone scans vary considerably in their size, contrast, brightness, and skeletal structure. In practice, many scans are only partial because only a determined part of the body is observed or due to the scanning limitations when the patient is too obese to fit in the screening area.

The idea of detecting the reference points is to find an image region that can easily be found in most cases. This region's reference point is then used as a starting location for the further reference point detection. Further reference point search is guided with the background knowledge of the spatial relations between bone regions and specific bones. The following image processing algorithms aid the search within some predefined boundaries (regions of interest, or ROIs): beam search, dynamic binarization, dilatation, skeletonization, and circles and ellipses fitting using LSM. Afterwards, all detected reference points are shifted vertically and horizontally to the neighboring regions with the highest uptake (intensity).

In our study, we observed that on only 2 images out of 467, the shoulders were not visible. Many other possible starting points are missing on the images more often (i.e., head, arms, one or both legs). Therefore, we chose shoulders as the main bone region to start with. The last assumption is the upward orientation of the image. This assumption is not limiting, since all scintigraphies are made with the same upward orientation.

In order to make the detection of reference points faster (linear time complexity) and more reliable, we tried to automatically detect peaks that would roughly cover the reference points and build some structural skeleton for reference point search guidance. The idea is similar to the SIFT [10] algorithm, where image features are also represented by the most outstanding pixels over different resolutions.

12.2.3.1 Detection of the Reference Point Candidates (Peaks)

There are numerous methods for detecting image peaks. We tried several. The SIFT [10] method returned too few peaks to guarantee a robust region detection on a vast variety of skeletons. More classic image filters, such as Canny edge detection, returned too many peaks. Peaks with such filters can be acquired by connecting the peaks to polylines, which are then reduced to vertices. We tried another customized approach based on orthogonal two-way Gaussian and linear filtering that mimics the SIFT algorithm. The algorithm smooths the image with the averaging window, whose size was experimentally determined. The algorithm works as shown in the pseudocode (see Algorithm 12.1).

AQ:1

Algorithm 12.1 Detection of relevant peaks Ω.

Require: Original image $I[x_i, y_i], i \in [0, m-1], j \in [0, n-1]$; Gaussian filter $G_{3\times 3} = \left(\frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \right)$; size of averaging window $W(w \times h$ where w and h are odd); d – minimal distance between detected peaks;

$$W_{w\times h} = \left(\frac{2}{w+h} \begin{bmatrix} 1 & 2 & \cdots & \frac{w+1}{2} & \cdots & 2 & 1 \\ 2 & 3 & \cdots & \frac{w+1}{2}+1 & \cdots & 3 & 2 \\ \vdots & \vdots & & \vdots & & \vdots & \vdots \\ \frac{h+1}{2} & \frac{h+1}{2}+1 & \cdots & \frac{w+h}{2} & \cdots & \frac{h+1}{2}+1 & \frac{h+1}{2} \\ \vdots & \vdots & & \vdots & & \vdots & \vdots \\ 2 & 3 & \cdots & \frac{w+1}{2}+1 & \cdots & 3 & 2 \\ 1 & 2 & \cdots & \frac{w+1}{2} & \cdots & 2 & 1 \end{bmatrix} \right)$$

Ensure: Relevant set of peaks Ω

1: $I_1 \Leftarrow (I \times G\{\text{Gaussian filter applied}\})$
2: $I_2 \Leftarrow (I_1 \times W_{w\times h}\{\text{bigger linear averaging filter applied}\})$
3: Add all elements from $I2$ to Ω_1 where both left and right side elements in I_2 have lower intensities {horizontal pass}
4: Add all elements from $I2$ to Ω_1 where both upper and lower elements in $I2$ have lower intensities {vertical pass}
5: Sort Ω_1 by element intensities
6: Add consecutively all elements I_i from Ω_1 to Ω where all elements in Ω have distance to $Ii > d$

The limited level of radioactivity injected due to radiation protection regulation in typical studies causes low intensities (count levels of gamma rays), which causes vague images of bone scans. Bone edges are better visible after images are filtered with some averaging algorithm (i.e., wavelet-based, median,

or Gaussian filter) [12]. A Gaussian filter is used in order to enhance the detection of peaks. The image is smoothed with the Gaussian filter where the pixels of higher intensities (> 90 gamma rays per pixel) are set to the predefined upper limit (90). Images smoothed in such a way are less prone to be obstructed by high-intensity lesions or artifacts, since at this stage we are interested only in finding anatomical reference points and not possible lesions.

In the case of scintigraphic images, we may introduce scale-variant methods, since the nature of this image modality guarantees a fixed scale (100 pixels $\approx 24\,\text{cm}$). The size of the averaging window $\text{Ww} \times h$ was experimentally set to $11 \times 5(2.6\,\text{cm} \times 1.2\,\text{cm})$ because it gave best results by the means of segmentation accuracy.

The reference points 1.1 are searched for using the detected peaks (Figure 12.3). Both images, anterior and posterior, are simultaneously processed in the

FIGURE 12.3: Example of detected image peaks on anterior and posterior image using Algorithm 12.1 ($d = 1, w \times h = 11 \times 5$) (all peaks are mirrored to the counter image).

same detection order. Detected points from the anterior image are mirrored to the posterior, and vice versa. Mirroring is possible because both images are taken at the same time and are therefore aligned and have the same size. Some point $pt_{(x,y)}$ on one image is mirrored to the other one as $pt'_{(x',y')}$, where $x' = image_{width} - x, y' = y$ on the other image. Some bones are better visible on the anterior and some on the posterior images due to the varying distances from the collimators (gamma-ray sensors). In each step, the detected reference points visible in both images are compared. The algorithm chooses the one that is estimated to better represent the desired reference point. For each reference point type (knee, ankle, ilium, pubis, shoulders, etc.), we determined the rules that decide better choice.

When deciding the better reference point of the anterior and posterior scan, two aspects are observed. The expected neighboring uptake and relations to other detected reference points on the basis of expected skeletal ratios (e.g., the length of the upper arm humerus is expected to be approximately 0.67 times the length of the spine, spanning from the neck nape to the sacrum bone). More detailed algorithms (source code) can be found in Šajn [17].

AQ:2

Bigger bone regions have the same shape and position in both images, but usually the edge of one side of the bone is better expressed on one image due to the distance to collimator, whereas the other side is better expressed on the other image. With combining peaks from both anterior and posterior images, both sides are clearly expressed (i.e., ilium bone). Some bone regions (skull, ilium, pubic bones, etc.) can be represented by some basic geometric shapes (circles, lines, and ellipses), which can be determined by using the LSM method. The fitting of the geometric shapes is improved also using the mirrored points from anterior to posterior image, and vice versa.

The order in which the reference points were detected was determined using the knowledge of the human anatomy as well as the physicians' recommendations. The anatomical knowledge is represented as a list of parameterized rules specific to each bone region. Rule parameters (e.g., thresholds, spatial and intensity ratios) were initially set by physicians and further refined on a separate tuning set.

12.2.3.2 Scintigraphy Segmentation

Respective skeletal regions are processed in the following order: shoulders, pelvis, head, thorax, and extremities. Here we briefly describe the idea of locating the bone regions. More detailed procedures are described in our previous work [18].

The two shoulder reference points as the main starting points are detected with the first peaks (see Figure 12.3) found from diagonal directions (Figure 12.4).

The pelvis is located at the end of the spine and has approximately the same width as the shoulders. In order to find the pelvis, the estimation of the spine position is required. This is performed with a beam search. The most

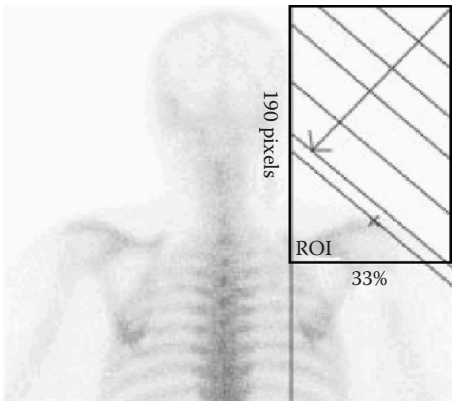


FIGURE 12.4: Locating shoulders.

identifiable bone in pelvic region is the ilium bone, which has higher uptake values than its neighboring soft tissue. The ilium bone has a circular shape in the upper part and is therefore convenient for circle fitting with LSM.

The pubis bone is detected by estimating the pubis ROI using the detected ilium location, the distance between detected ilium circle centers, and the angle between them. The experimentally determined ROI's size is narrowed with the binarization algorithm. Circles representing the two ilium and pubic bones are detected with LSM using the detected peaks. AQ:3

When at least the image orientation and the location of the shoulders are known, some part of the neck or the head is visible between the shoulders. Finding the head is not difficult, but determining its orientation is, especially in cases where a part of the head is not visible in the scan. The most reliable method for determining the head orientation and position is ellipse fitting of the head contour, as shown in the Figure 12.5.

Vertebrae have more or less specific intervertebral spaces [19]; the only problem is that on a bone scintigram, only a planar projection of the spine is visible. Since the spine is longitudinally curved, the vertebral spatial relations vary due to different axial orientation of the patients. The average intervertebral disc spaces have been experimentally determined from normal skeletons. After the approximate spine ROI is determined, it is divided into 10 vertical parts according to the predefined intervertebral disc.

Ribs are the most difficult skeletal region to detect because they are quite unexpressive on bone scans, their formation can vary considerably, and their contours can be disconnected in the case of stronger pathology. Using the detected vertebrae, the ribs' ROIs are defined. For the rib contour detection, we use the morphology-based image operations, particularly three well-known algorithms: dynamic binarization, dilation/skeletonization [20], and Hough transform [21].

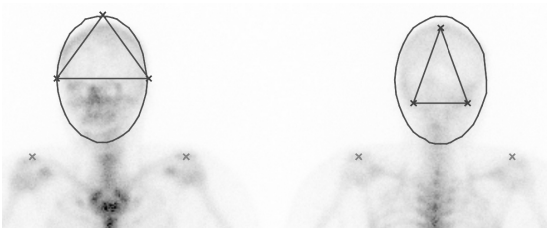


FIGURE 12.5: Head detection with ellipse fitting sizes.

Extremities are often partly absent from the whole-body scan because of the limited gamma camera detector width. The regions of humerus, ulna, and radius as well as femur, tibia, and fibula bones are located with the use of beam search. The detection is designed in such a way that a part or the entire extremity or head may not be visible in the scan.

12.3 The ArTeX Algorithm

12.3.1 Related Work

Researchers have tried to characterize textures in many different ways. Most texture features are based on structural, statistical, or spectral properties of the image. Some methods use textural features that include several of these properties. Well-known statistical features are based on gray-level cooc-
AQ:4 currence statistics [22], which are used in the Image Processor program [23]. Examples of structural features are features of Voronoi tesselation [24], representations using graphs [25], representations using grammars [26], and representations using association rules [27]. The spectral features are calculated in space, which is closely related to textural features—for example, frequency and amplitude. The most frequently used space transformations are the Fourier, Laws [28], Gabor [29], and wavelet transforms.

12.3.2 Association Rules

Association rules were introduced by Agrawal, Imielinski, and Swami [30] back in 1993. The following is a formal statement of the problem: Let $\mathcal{I} = \{i_1, i_2, \ldots, i_m\}$ be a set of literals, called items. Let $\mathcal{D}$ be a set of transactions, where each transaction T is a set of items such that $T \subseteq \mathcal{I}$. We say that a transaction T contains X if $X \subseteq T$. An *association rule* is an implication of the form $X \Longrightarrow Y$, where $X \subset \mathcal{I}, Y \subset \mathcal{I}$ and $X \cap Y = 0$. The rule $X \Longrightarrow Y$ holds in the transaction set $\mathcal{D}$ with confidence c if $c\%$ of transactions in $\mathcal{D}$

that contain X also contain Y. The rule $X \Longrightarrow Y$ has support s in the transaction set $\mathcal{D}$ if $s\%$ of transactions in $\mathcal{D}$ contain $X \cup Y$. The problem of discovering association rules says: Find all association rules in transaction set $\mathcal{D}$ with confidence of at least *minconf* and support of at least *minsup*, where *minconf* and *minsup* represent the lower boundary for confidence and support of association rules.

12.3.3 Texture Representation

The use of association rules for texture description was independently introduced by Rushing and others [27]. Here we present a slightly different approach, which uses different texture representation and a different algorithm for association rules induction and which we developed before we become aware of the work by Rushing and colleagues.

Association rules are most widely used for data mining of very large relational databases. In this section, we give a representation of texture, which is suitable for processing with the association rules algorithms. To apply the association rules algorithms on textures, we must first define the terms used for association rules in the context of textures.

Pixel *A* of a texture P is a vector $\vec{A} = (X, Y, I\,) \in P$, where X and Y represent the absolute coordinates and I represents the intensity of pixel A.

Root pixel $\vec{K}$ is the current pixel of a texture $\vec{K} = (X_K, Y_K, I_K)$.

R neighborhood $N_{R,\vec{K}}$ is a set of pixels located in the circular area of radius R with root pixel $\vec{K}$ at the center. Root pixel $\vec{K}$ itself is not a member of its neighborhood.

$$N_{R,\vec{K}} = \{(X, Y, I) | \delta \leq R\} \setminus \vec{K}$$

$$\delta = \left[\sqrt{(X_K - X)^2 + (Y_K - Y)^2} + 0.5 \right] \tag{12.1}$$

Transaction is a set of elements based on its corresponding neighborhood. The elements of transaction are represented with Euclidean distance and intensity difference from the root pixel.

$$T_{R,\vec{K}} = \left\{ (\delta, I_K - I) | (X, Y, I) \in N_{R,\vec{K}} \right\}$$

$$\delta = \left[\sqrt{(X_K - X)^2 + (Y_K - Y)^2} + 0.5 \right] \tag{12.2}$$

Transaction element is a two-dimensional $(r, i) \in T_{R,\vec{K}}$ where the first component represents the Euclidean distance from the root pixel

and the second component represents the intensity difference from the root pixel.

Association rule is composed of transaction elements; therefore, it looks like this:

$$(r_i, i_1) \wedge \cdots \wedge (r_m, i_m) \Longrightarrow$$
$$(r_{m+1}, i_{m+1}) \wedge \cdots \wedge (r_{m+n}, i_{m+n})$$

Transaction set ${}^{D}P$, R is composed of transactions that are derived from all possible root pixels of a texture P at certain neighborhood size R:

$$D_{P,R} = \left\{ T_{R,\vec{K}} | \forall \vec{K} : \vec{K} \in P \right\}$$

This representation of a texture replaces the exact information of location and intensity of the neighboring pixels with more indecisive information of the distance and the relative intensity of neighboring pixels. This description is also rotation invariant.

Figure 12.6 illustrates the association rule $(1,1) \wedge (2,10) \Longrightarrow (1,15) \wedge (3,5)$, which can be read as follows: If a pixel of intensity 1 is found on distance 1 and a pixel of intensity 10 is found on distance 2, then there is also a pixel of intensity 15 on distance 1 and a pixel of intensity 5 on distance 3.

This representation is almost suitable for processing with general association rule algorithms. What is still to be considered is the form of a transaction element. Association rule algorithms expect scalar values for transaction elements, whereas our representation produces a two-dimensional vector for a transaction element. Luckily, this issue can be easily solved. Let us say that the intensity of each texture point can have values from interval $[0..(Q-1)]$

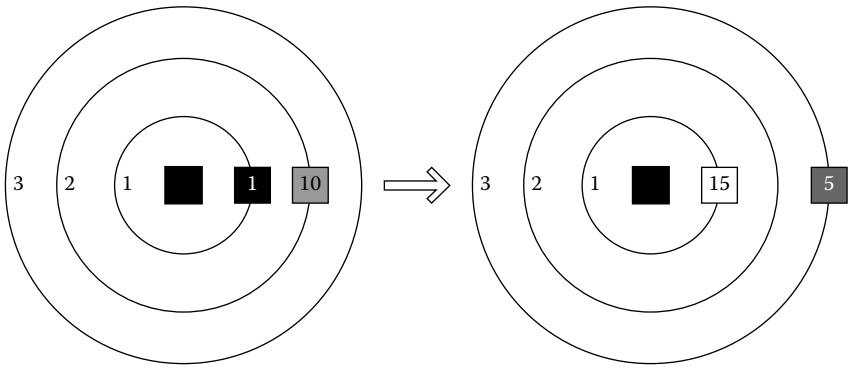


FIGURE 12.6: An illustration of association rule $(1,1) \wedge (2.10) \Rightarrow (1,15) \wedge (3,5)$.

and that the neighborhood size is R. Take some transaction element (r, i), where i has a value from $[-(Q-1)..+(Q-1)]$ and r has a value from $[1..R]$. What is needed here is a bijective mapping that transforms each vector to its scalar representation. This can be achieved in many ways. A possible and quite straightforward solution is:

$$s = (2Q-1)(r-1) + i + (Q-1)$$

The transformation is also reversible:

$$r = 1 + s \operatorname{div}(2Q-1)$$
$$i = s \bmod (2Q-1) - (Q-1)$$

Now it is possible to define a transaction that suits the general association rule algorithms:

$$T_{R,Q,\vec{K}} = \left\{ s \;\middle|\; \begin{array}{l} (r,i) \in T_{R,\vec{K}} \\ s = (2Q-1)(r-1) + i + (Q-1) \end{array} \right\}$$

And finally, we obtain the appropriate transaction set definition:

$$D_{P,R} = \left\{ T_{R,Q} | \forall \vec{K} : \vec{K} \in P \right\}$$

12.3.4 From Association Rules to Feature Description

Using association rules on textures allows us to extract a set of features (attributes) for a particular domain of textures. The ArTeX algorithm is defined with the following steps:

- *Select a (small) subset of images* F *for feature extraction.* The subset F can be considerably small. Use at least one example of each typical image in the domain—that is, at least one sample per class, or more if the class consists of subclasses.

- *Preprocessing of images in* F. Preprocessing involves the transformation of images to grayscale if necessary, the quantization of gray levels, and the selection of proper neighborhood size R. The initial number of gray levels per pixel is usually 256. The quantization process downscales it to, say, 32 levels per pixel. Typical neighborhood sizes are 3, 4, 5.

- *Generate association rules from images in* F. Because of the representation of texture, it is possible to use any algorithm for association rules extraction. We use *Apriori* and *GenRules*, as described in [30].

- *Use generated association rules to extract a set of features.* There are two features associated with each association rule: support and confidence. Use these two attributes of all association rules to construct a feature set. The number of extracted features is twice the number of association rules (which could be quite a lot).

To clarify what we said, we also provide a formal algorithm (see Algorithm 12.2). The algorithm takes five input parameters: a set of images I, neighborhood size R, texture quantization Q, minimum support *minsup*, and minimum confidence *minconf*. Functions φ_{sup} and φ_{conf} are used to calculate support and confidence given an image and an association rule. The output of the algorithm is feature set matrix d, where $d_{i,j}$ represents jth feature of image i.

Algorithm 12.2 ArTeX (images I, radius R, quantization Q, *minsup*, *minconf*).

Select F so that $F \subset I$
Preprocess(F, R, Q)
for all $f \in F$ **do**
  $\mathcal{D}$ = transactionModel(f, R, Q) STATE r_1 = apriori$(\mathcal{D}, \text{minsup})$
  r_2 = genRules$(r_1, \mathcal{D}, \text{minconf})$
  $\rho_{sup} = \rho_{sup} \cup r_1$ {itemsets with support > minsup}
  $\rho_{conf} = \rho_{conf} \cup r_2$ {rules with confidence > minconf}
end for
$i = 0$
for all $f \in (I \setminus F)$ **do**
  $j = 0$
  for all $\varrho \in \rho_{sup}$ **do**
    $d_{i,j} = \varphi_{sup}(f, \varrho)$ {jth attribute of ith image}
    $j = j + 1$
  end for
  for all $\varrho \in \rho_{conf}$ **do**
    $d_{i,j} = \varphi_{conf}(f; \varrho)$
    $j = j + 1$
  end for
  $i = i + 1$
end for
return d {d is a matrix of attribute values}

12.3.5 Extending the Parameter Set

Our model of texture is such that the structure of the association rule also describes some aspects of the textural structure. Since we are interested in the parametric description of a texture, this structure has to be represented with one or more parameters. Until now, we presented the basic algorithm, which uses only basic interestingness measures, support, and confidence, which were defined together with association rules [30]. They are still most widely used today, but there are some concerns, especially with confidence measure, which can be misleading in many practical situations, as shown by Brin and others [31]. They also offered an alternative to evaluate association rules using

χ^2 test. Contrary to confidence measure, χ^2 test could be used to find both positively and negatively correlated association patterns. However, the χ^2 test alone may not be the ultimate solution because χ^2 test does not indicate the strength of correlation between items of association pattern. It only decides whether items of association pattern are independent of each other; thus, it cannot be used for ranking purposes.

We use χ^2 test just to select interesting association patterns, which are later described by the Pearson correlation coefficient (ϕ-coefficient) as advised in Tan and Kumar [32]. ArTeX also uses an additional interestingness measure, which was selected with thorough experiments on various domains from a subset of collections made by Tan and colleagues [33]. From all tested measures, the J-measure gave best results [34]. J-measure for rule $A \rightarrow B$ is defined as follows:

$$P(A;B)\ P(A,B)\log\left(\frac{P(B|A}{P(B)}\right) + P(A,\bar{B})\log\left(\frac{P(\bar{B}|A}{P(\bar{B})}\right)$$

12.4 Multiresolutional Parameterization

Why use more resolutions? Digital images are stored in matrix form, and algorithms for pattern parameterization basically use some relations between image pixels (usually first- or second-order statistics). By using only a single resolution, we may miss the big picture and proverbially not see the forest for the trees. Since it is too computationally complex to observe all possible relations between at least any two pixels in the image, we have to limit the search to some predefined neighborhood. These limitations make relations vary considerably over different resolutions. This means that we may get completely different image parameterization attributes for the same image at different scales.

12.4.1 Parameters from Many Resolutions

In different existing multiresolutional approaches [35–37], many authors are using only more resolutions, which are not determined on the basis of image contents. Usually two or three resolutions are used. Authors report better classification results when using more resolutions and also observe that when using more than three resolutions, the classification accuracy starts to deteriorate. We have observed that in many cases authors use a set of resolutions by exponentially decreasing the resolution size ($\frac{100i}{n}, i = 0..n-1$). However, we noticed that in many cases, equidistant selection of resolutions ($\frac{100i}{n}, i = 1..n$, where n is the number of resolutions used) gives better results. When using exponential forms of resolutions, less pattern content is examined and consecutively less significant attributes are derived. Another frequently

AQ:5

used multiresolutional approach is the wavelet transform [38], which describes textures with measures calculated with iterative image division. None of the procedures mentioned above observe the contents of images.

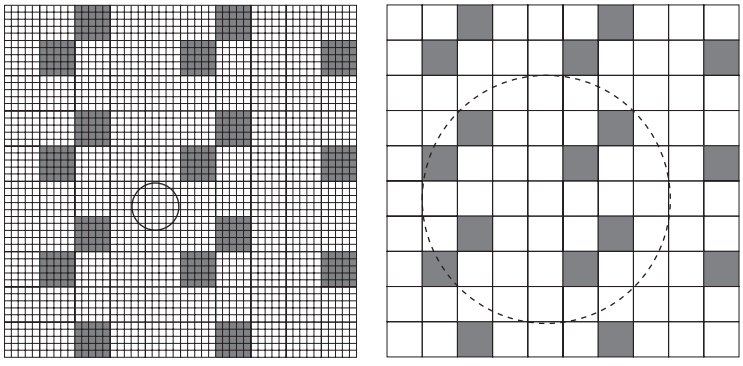
Another extension of parameters for texture parameterization comes from the issue of pattern's scale. Not every combination of scale and neighborhood size can guarantee that the pattern is detected. The problem is illustrated in Figure 12.7.

To increase the possibility that the pattern will be detected, we propose a framework in which the extraction of attributes is repeated at different texture resolutions and combined in one feature vector.

12.4.2 Automatic Selection of a Small Subset of Relevant Resolutions

The idea for the algorithm for automatic selection of a small subset of relevant resolutions is derived from the well-known SIFT algorithm [10]. The SIFT algorithm is designed as a stable local feature detector represented as a fundamental component of many image registration and object recognition algorithms. Since we are not interested in detecting stable image key points but in detecting resolutions at which the observed image has most extremes, we devised a new algorithm, ARes (see Algorithm 12.3) for determining the resolutions for which more informative features can be obtained. The algorithm was designed especially for the ArTeX parameterization algorithm (see Algorithm 12.2) but also usually improves the results with other parameterization algorithms, as can be seen in Section 12.5. Resolutions for the bone scintigraphy are also determined using ARes algorithm.

AQ:6



(a) A Very inadequate resolution (b) An adequaten resolution

FIGURE 12.7: Detecting patterns at different scales.

The ARes algorithm consequently resizes the image from 100% down to some predefined lowest threshold at some fixed step when detecting the appropriate resolutions. Both the lowest threshold and the resolution step are determined using the observed image data set. At each resize step, the peaks are counted. Peaks are represented by pixels that differ from their neighborhood either as highest or lowest intensity. This algorithm can also be implemented with the difference-of-Gaussian (DOG) [10] method, which improves the time complexity with a lower number of resizes required to search the entire resolution space. AQ9

The detected peak counts are recorded over all resolutions as a histogram. From the histogram, the best resolutions are detected as the highest counts. The number of resolutions we want to use in our parameterization is predefined by the user. When there are several equal counts, we choose resolutions as diverse as possible.

Algorithm 12.3 Algorithm ARes for detecting a small subset of relevant resolutions.

Require: Set of input images Θ with known classes, number of desired resolutions η, number of images to inspect in each class γ, radius ϕ used by the parameterization algorithm later in the process

Ensure: Subset of resolutions

1: $$W_{\max} = \max_{i=1}^{|\Theta|}(\Theta i_{(width)}), H_{max} = \max_{i=1}^{|\Theta|}(\Theta i_{(height)})$$
{find the biggest image height and width}

2: extend the image sizes $\Theta_i \in \Theta$ to $W_{max} \times H_{max}$ by adding a frame of intensity equal to the average intensity of the original image Θ_i. New resized images are saved in the set Θ' {image sizes must be unified in order to be able to compare resolutions over different images}

3: $\delta = \frac{2*\phi}{3} \cdot \frac{1}{\{W_{max} \times H_{max}\}}$ {set the resize step}

4: for each class, add γ randomly selected images from the set Θ' into the set Θ_1

5: $\Omega = \{\}$

6: **for** $(\forall \theta \in \Theta_1)$ **do**

7: $\nu = 1.0$ {start with 100% resolution}

8: **while** $(\min W_{max}, H_{max} \cdot \nu > 3 \cdot \phi)$ **do**

9: $\theta_1 = resize(\theta, \nu)$ {change the observed image's size}

10: Find local peaks in θ_1 by comparing each pixel's neighborhood inside $[3 \times 3]$ window

11: Add the pair {ν, number of peaks} into the set Ω

12: $\nu = \nu - \delta\nu$

13: **end while**

14: **end for**

15: order the set Ω by the number of descending peaks and resolutions

16: add first η resolutions from the ordered set Ω into the final set

12.5 Results in Diagnostics of the Whole-Body Scintigrams

12.5.1 Segmentation

Approximately half of the scans were used for tuning the parameters required for reference point detection and the other half to test the segmentation process. All 246 patients examined between October 2003 and March 2004 were used as the tuning set, and the 221 patients examined between April and June 2004 were used as the test set. In the tuning set, 38.9% of the images showed various nonosseous uptakes, 47.5% images contained visible injection points, and 6.8% were of children/adolescents with visible growth zones. Similar distribution was found in the test set (34.5% nonosseous uptakes, 41.0% visible injection points, and 2.85% children/adolescents). Most of the artifacts were minor radioactivity points from urine contamination in the pelvic region or other parts (81.4% of all artifacts), whereas only few other types were observed (urinary catheters 13%, artificial hips 4%, and lead accessories 1.6%). We observed that there were no incorrectly detected reference points in children/adolescents with the visible growth zones, since all the bones are homogenous, have good visibility, and are clearly divided with growth zones. The segmentation algorithm works for adult and children/adolescents patients without any user intervention. Results of detecting the reference points on the test set are shown in Table 12.1.

Results are presented for different types of pathologies because we expected the degree of pathology to affect the quality of the detection process. The results show that there are no major differences in the reference point detection between different pathological groups.

The algorithm was implemented in a system called Skeleton 1.2 [17] (written in Java 2 SE, version 1.4.2) (see Figure 12.8). The current system includes image editing, filtering with different linear filters, conversions to other formats, storing images in the database with batch procedures, XML exports of reference points, manual correction of reference points, image annotation, and region localization using the scalable correlation. We also tested the accuracy of the bone region localization using correlation with generalized bone region images. In addition to bigger time complexity, the latter approach gave worse results compared to our segmentation algorithm. Therefore, all high-complexity algorithms were avoided, and applied algorithms were optimized in the sense of computational complexity. The detection of reference points on both anterior and posterior images takes approximately 3 seconds on a Pentium 4 PC with 2.8 GHz, 1 GB RAM.

12.5.2 Diagnosing Pathologies with Machine Learning

The obtained reference points can be used in two ways to segment a scan. One way is to segment the scintigram only by extracting bones along the

T&F Cat # C6059 C6059_C012 Page: 369 2008-11-4

TABLE 12.1: Number of correctly detected reference points on the test set in different degrees of pathologies. Both frequencies and percentages are given.

Bone	**No Pathology**		**Slight Pathology**		**Strong Pathology**		**Super-Scan**		**All**	
Females	**22**		**78**		**15**		**1**		**116**	
Ilium	22	100%	76	97%	13	87%	0	0%	111	96%
Pubis	22	100%	76	97%	14	93%	1	100%	113	97%
Trochanter	22	100%	77	99%	15	100%	1	100%	115	99%
Shoulder	22	100%	78	100%	14	93%	1	100%	115	99%
Extremities	21	96%	74	95%	15	100%	1	100%	111	96%
Spine	22	100%	78	100%	14	93%	1	100%	115	99%
Ribs	16	73%	70	90%	12	80%	1	100%	99	85%
Neck	22	100%	76	97%	15	100%	1	100%	114	98%
Average	96%		97%		93%		88%		96%	
Males	24		55		24		2		105	
Ilium	24	100%	55	100%	20	83%	2	100%	101	96%
Pubis	22	92%	54	98%	23	96%	2	100%	101	96%
Trochanter	24	100%	55	100%	24	100%	2	100%	105	100%
Shoulder	24	100%	55	100%	24	100%	2	100%	105	100%
Extremities	20	83%	48	87%	24	100%	2	100%	94	90%
Spine	24	100%	53	96%	24	100%	2	100%	103	98%
Ribs	19	79%	46	84%	24	100%	2	100%	91	87%
Neck	22	92%	53	96%	24	100%	2	100%	101	96%
Average	93%		95%		97%		100%		95%	
Sum	46		133		39		3		221	
Ilium	46	100%	131	98%	33	85%	2	67%	212	96%
Pubis	44	96%	130	98%	37	95%	3	100%	214	97%
Trochanter	46	100%	132	99%	39	100%	3	100%	220	100%
Shoulder	46	100%	133	100%	38	97%	3	100%	220	100%
Extremities	41	89%	122	92%	39	100%	3	100%	205	93%
Spine	46	100%	131	98%	38	97%	3	100%	218	99%
Ribs	35	76%	116	87%	36	92%	3	100%	190	86%
Neck	44	96%	129	97%	39	100%	3	100%	215	97%
Average	95%		94%		96%		96%		96%	

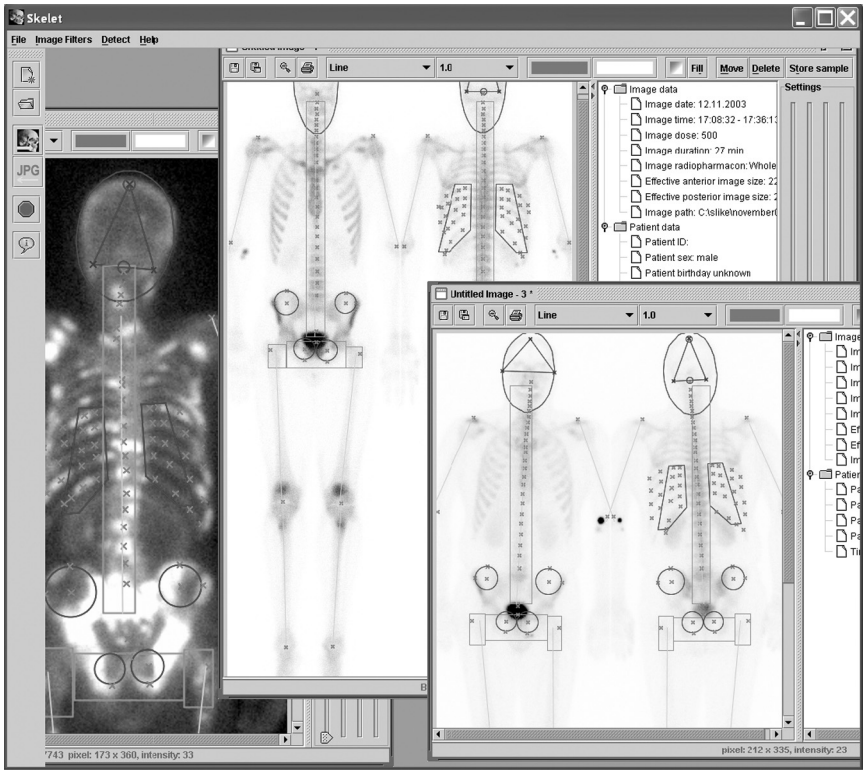


FIGURE 12.8: (See color insert following page xxx.) View of the program Skeleton.

detected reference points with some predefined offset width. This type of segmentation was used in the classification part of our study. Since there is only a limited number of reference points, we can extract only 26 bones or bone regions (i.e., extremities, lumbal, thoracic, and cervical spine; 10 ribs; bones in pelvic region; and the head). Some extracted bones are shown in Figure 12.9(c).

Another possibility of segmenting the scintigram is to map some standard, predefined skeletal mask over the scan observed using the detected reference points, as shown in Figure 12.9(b). With this approach, we also need some virtual reference points, which are defined with the detected reference points; see Figure 12.9(a). The skeletal mask represented with polygons can be defined by the radiologist on some scans. The algorithm triangulates the scan using the detected reference points. A new scan is triangulated in the same way so the defined skeletal mask can be mapped. Mapping is carried out with the linear transform for each triangle. The defined skeletal mask polygons

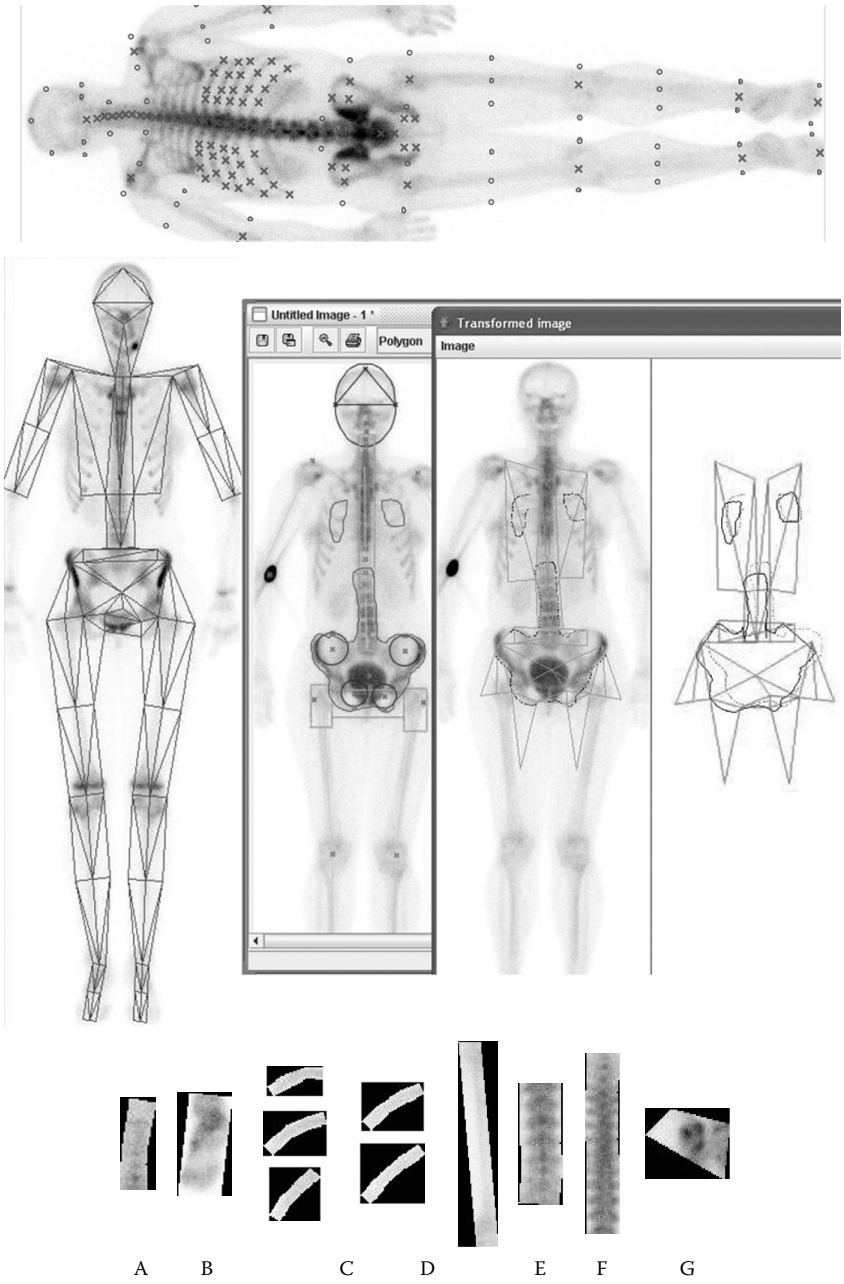


FIGURE 12.9: Using reference points for scintigraphy segmentation. (a) Additional virtual reference points (b) Triangulation mask and an example of a mapped annotation between different scans (c) Examples of extracted bones (A, cervical spine; B, foot; C, ribs; D, femur; E, lumbal spine; F, thoracic spine; G, sacroiliac joint).

are transformed with a linear operation that multiplies the polygon vertices $p_{im}(x_{im}, y_{im})$ with the matrix A calculated from the belonging triangle coordinates in the original image $v_{1o}(x_{1o}, y_{1o})$, $v_{2o}(x_{2o}, y_{2o})$, $v_{3o}(x_{3o}, y_{3o})$ and in the new image $v_{1m}(x_{1m}, y_{1m})$, $v_{2m}(x_{2m}, y_{2m})$, $v_{3m}(x_{3m}, y_{3m})$.

$$A = \begin{bmatrix} x_{1m} & x_{2m} & x_3m \\ y_{1m} & y_{2m} & y_3m \\ 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_{1o} & x_{2o} & x_{3o} \\ y_{1o} & y_{2o} & y_{3o} \\ 1 & 1 & 1 \end{bmatrix}^{-1}, P_{im} = A \cdot P_{io}$$

When all reference points are obtained, every bone is assigned a portion of original scintigraphic image according to relevant reference points.

The results of our parameterization with our ArTeX algorithm are compared with the results of four other image parameterization algorithms (Haar wavelets [38], Laws filters [39], Gabor filters [40], image processor [23] (implements many parameters of the first- and second-order statistics [22]), and Laws texture measures [39]).

The images of bones were described with several hundreds of automatically generated attributes with the ArTeX algorithm. Attributes are invariant to rotation and illumination changes. Rotation invariance is very important in this case because it compensates for different patients' positions inside the camera, whereas the illumination compensates for different absorptions of radiopharmaceutical throughout the body. Attributes were used for training the naive Bayesian classifier implemented in Weka [41]. In our preliminary experiments, various pathologies were not discriminated; that is, bones were labeled with only two possible diagnoses: no pathology or pathology.

From our complete set of 467 patients, pathologies were thoroughly evaluated by physicians only for 268 patients. These 268 patients were used for the experiments with machine learning. In 21% of scans, no pathology or other artifacts were detected by the expert physicians. In the remaining 79% of the scans, at least one pathology or artifact was observed. All bones were classified as normal or pathological and grouped into 10 regions. For each region, we have a binary classification problem. For machine learning, we used all pathological bones and randomly sampled healthy bones so that in all 10 problems, 30% of the bones were pathological, and 70% were healthy. Results were evaluated with tenfold cross validation and are shown in Tables 12.2 and 12.4. Table 12.2 gives the classification accuracy with the ArTeX algorithm, and Table 12.4 compares the other algorithms. The best results were achieved when ArTeX with J-measure and resolutions determined with ARes were used for parameterization of images. The Friedman rank test showed that ArTeX was significantly better ($\alpha = 0.05$) than the Laws and Haar algorithms; however, the difference compared with the Gabor and image processor algorithms using resolutions determined with ARes is not significant.

Table 12.3 gives sensitivity and specificity of the classification results. ArTeX achieved the best specificity. The Friedman rank test showed that

TABLE 12.2: Experimental results (classification accuracy [%]) with machine learning on two-class problem at different resolutions. Parameterized with ArTeX using the J-measure.

Bone Region	100% Resolution	4 Equidistant Resolutions	4 Resolutions with ARes
Cervical spine	76,00	72,00	75,00
Feet	84,44	84,44	85,56
Skull posterior	70,00	80,00	80,00
Ilium bone	88,48	90,30	90,30
Lumbal spine	69,10	74,42	71,41
Femur and tibia	84,15	84,15	86,41
Pelvic region	88,57	91,43	94,29
Ribs	94,31	95,42	95,42
Scapula	95,00	95,00	95,00
Thoracic spine	71,67	71,67	81,12
Average	82,17	83,88	85,45

TABLE 12.3: Specificity and sensitivity of the classification.

Bone Region	100% Resolution		Equidistant		ARes	
	Sp. %	Sen. %	Sp. %	Sen. %	Sp. %	Sen. %
Cervical spine	88,89%	42,86%	88,89%	28,57%	86,10%	56,08%
Feet	93,75%	61,54%	93,75%	61,54%	93,75%	65,38%
Skull posterior	84,62%	0,00%	100,00%	0,00%	100,00%	0,00%
Ilium bone	91,36%	81,82%	93,83%	81,82%	93,83%	81,82%
Lumbal spine	80,68%	43,90%	88,64%	43,90%	81,82%	48,78%
Femur and tibia	91,20%	67,31%	95,20%	57,69%	94,40%	73,08%
Pelvic region	94,00%	75,00%	96,00%	80,00%	98,00%	85,00%
Ribs	95,16%	92,00%	98,39%	88,00%	98,39%	88,00%
Scapula	100,00%	81,82%	100,00%	81,82%	100,00%	81,82%
Thoracic spine	82,93%	43,75%	85,37%	37,50%	85,37%	61,20%
Average	90,26%	59,00%	94,01%	56,08%	93,16%	64,12%

TABLE 12.4: Experimental results with other algorithms.

Algorithm	100% Resolution	Equidistant	ARes
ArTeX (no J-measure)	83,25	83,00	84,91
Gabor	83,61	84,41	84,12
Haar	80,16	79,80	81,82
Laws	82,60	83,00	83,39
Image processor	81,47	83,59	84,46

in the specificity of the bone classification, ArTeX was significantly better ($\alpha = 0.05$) than the Haar and Gabor algorithms; however, the difference compared with the image processor and Laws algorithms is not significant. Sensitivity was not significantly different in any comparison of ArTeX to other algorithms.

12.6 Discussion

The detection of the reference points gave excellent results for all bone regions except for the ribs and the extremities, where some parts were missing in the scan. The extremities were mainly misdetected in the cases where the humerus bone was partially missing.

As expected, the detection of ribs was the most difficult. The results show that in 14% to 20% of scans, there were difficulties in detecting ribs. The ribs in the thoracic areas were hard to follow due to a vague expression. Generally, one rib was missed or not followed to the very end. We intend to improve this problem in the future. In the present system (Figure 12.8), such reference points can be manually repositioned by the expert physician.

Since a robust segmentation algorithm should not fail on partial skeletal images, which is often the case in clinical routine (18% of the scans in our study), special attention was paid to such cases (e.g., amputees and skeletal parts entirely invisible in the scan). In our results, such cases do not stand out from the normal scans.

The automatically detected reference points can be used for mapping a standard skeletal reference mask, which we believe is the best way to find individual bones on scintigrams, since individual bones are often not expressive enough that their contours can be followed. Examples of such mask mapping and extracted bones are shown on Figure 12.9(b) and 12.9(c).

Segmented bone images can be treated as textures because they have similar structure over a certain bone region. Bones as textures can be parameterized with some algorithms for pattern parameterization. Pathology represents some unusual texture patterns that can be distinguished from healthy cases. The ArTeX algorithm for pattern parameterization alone performs just as well as other algorithms in terms of classification accuracy. When we extend the parameter set with the J-measure and use the multiresolutional approach, the performance greatly improves. The resolutions for the multiresolutional approach must be carefully selected, as shown in the results. Resolution selection is performed by the algorithm ARes, which is also described in the chapter and improves ArTeX as well as other parameterization algorithms (e.g., Haar, Laws, and image processor). The same performance is achieved in other medical domains (e.g., ischemic heart disease), which are not discussed in this chapter.

12.6.1 Future Research Directions

The presented algorithms for pattern parameterization open a new research area of multiresolution image parameterization and enable many applications in medical, industrial, and other domains where textures or texturelike surfaces are classified. The ARes algorithm can be improved with additional resolution search refinements, which would be more domain oriented. In our case, we plan to study malignant pathologies and to seek different criteria for resolution quality evaluation.

While our experimental results with machine learning are quite satisfactory, they were obtained for a simplified (two class) problem only. Simply extending a problem to a multiclass paradigm is not acceptable, since the bone may be assigned several different pathologies at the same time. We are currently developing a new approach in which the problem is rephrased to the multilabel learning problem, and each bone will be labeled with a nonempty subset of all possible labels [42, 43].

12.7 Conclusions

The presented computer-aided system for bone scintigraphy is a step toward automating the routine medical procedures. Some standard image processing algorithms were tailored and used in combination to achieve the best reference point detection accuracy on scintigraphic images, which have technically very low resolution. Because of poorer image resolution compared to radiography, the presence of artifacts and pathologies necessitates that algorithms use as much background knowledge of anatomy and spatial relations of bones as possible in order to work satisfactorily. This combination gives quite good results, and we expect that further studies on automatic scintigram diagnosis using reference points for image segmentation will give more accurate and reliable results than presented in previous studies that did not use segmentation.

For improving classification accuracy in medical image domains, we encourage the multiresolutional parameterization approach. As we have observed, the resolutions should be determined according to the properties of the observed domain.

This approach opens a new view on automatic bone scintigraphy evaluation, since in addition to detection of pointlike high-uptake lesions, it also offers

- more accurate and reliable evaluation of bone symmetry when looking for skeletal abnormalities;

- detection of a greater number of abnormalities, since many abnormalities can be spotted only when the symmetry is observed (differences in length, girth, curvature etc.);
- detection of lesions with low-uptake or lower activity due to metallic implants (e.g., artificial hip);
- possibility of comparing uptake ratios among different bones;
- more complex pathology detection by combining pathologies of more bones (e.g., arthritis in joints);
- possibility of automatic reporting of bone pathologies in written language for educational purposes.

The machine learning approach described in this work is in an early stage of development. However, the preliminary results are encouraging, and the multiresolutional approach will make them more useful for clinical applications.

References

[1] V. Mäuller, J. Steinhagen, M. deWit, and H. K. Bohuslavizki. Bone scintigraphy in clinical routine. *Radiology & Oncology*, 35(1):21–30, 2001.

[2] A. Hendler and M. Hershkop. When to use bone scintigraphy. It can reveal things other studies cannot. *Postgraduate Medicine*, 104(5):54–66, 1998.

[3] M. Noguchi, H. Kikuchi, M. Ishibashi, and S. Noda. Percentage of the positive area of bone metastasis is an independent predictor of disease death in advanced prostate cancer. *British Journal of Cancer*, (88):195–201, 2003.

[4] J. Bernauer. *Zur semantischen rekonstruktion medizinischer begriffssysteme. habilitationsschrift.* Institut für Medizinische Informatik, Universitäat Hildesheim, 1995.

[5] K. C. Berning. *Zur Automatischen Befundung und Interpretation von Ganzkäorper-Skelettszintigrammen.* PhD thesis, Institut fäur Medizinische Informatik, Universitäat Hildesheim, 1996.

[6] A. Benneke. Konzeption und realisierung eines semi-automatischen befundungssystems in Java und an-bindung an ein formalisiertes begriffssystem am beispiel der skelettszintigraphie. Diplom arbeit, Institut fäur Medizinische Informatik, Universitäat Hildesheim, mentor Prof. Dr. D. P. Pretschner, 1997.

[7] M. N. Maisey, T. K. Natarajan, P. J. Hurley, and H. N. Jr. Wagner. Validation of a rapid computerized method of measuring 99mTc pertechnetate uptake for routine assessment of thyroid structure and function. *Journal of Clinical Endocrinology & Metabolism*, 36:317–322, 1973.

[8] T. K. Yin and N. T. Chiu. A computer-aided diagnosis for locating abnormalities in bone scintigraphy by a fuzzy system with a three-step minimization approach. *IEEE Transactions on Medical Imaging*, 23(5):639–654, 2004.

[9] M. Kukar, I. Kononenko, C. Grošeli, K. Kralj, and J. Fettich. Analysing and improving the diagnosis of ischaemic heart disease with machine learning. *Artificial Intelligence in Medicine*, 16:25–50, 1999.

[10] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.

[11] H. Murase and S. Nayar. Visual learning and recognition of 3-D objects from appearance. *International Journal of Computer Vision*, 14(1):5–24, 1 1995.

[12] G. Jammal and A. Bijaoui. DeQuant: A flexible multiresolution restoration framework. *Signal Processing*, 84(7):1049–1069, 2004.

[13] I. Kononenko and M. Kukar. *Machine Learning and Data Mining: Introduction to Principles and Algorithms*. Chichester, UK: Horwood, 2007.

[14] M. Kukar, L. Šajn, C. Grošelj, and J. Grošelj. Multi-resolution image parameterization in sequential diagnostics of coronary artery disease. In R. Bellazzi, A. Abu-Hanna, and J. Hunter, Eds., *Artificial Intelligence Image Segmentation and Parameterization for Automatic Diagnostics of Whole-Body Scintigrams in Medicine*, pp. 119–129. Berlin: Springer, 2007.

[15] I. Kononenko. Inductive and Bayesian learning in medical diagnosis. *Applied Artificial Intelligence*, 7:317–337, 1993.

[16] M. G. Weiner, L. Jenicke, V. Müller, and H. K. Bohuslavizki. Artifacts and non-osseous uptake in bone scintigraphy. Imaging reports of 20 cases. *Radiology & Oncology*, 35(3):185–91, 2001.

[17] L. Šajn. Program source code, 2007. lkm.fri.uni-lj.si/skelet/skelet.zip.

[18] L. Šajn, M. Kukar, I. Kononenko, and M. Milčinski. Computerized segmentation of whole-body bone scintigrams and its use in automated diagnostics. *Computer Methods & Programs in Biomedicine*, 80(1):47–55, 2005.

[19] G. Holmaas, D. Frederiksen, A. Ulvik, S. O. Vingsnes, G. Ostgaard, and H. Nordli. Identification of thoracic intervertebral spaces by means of surface anatomy: A magnetic resonance imaging study. *Acta Anaesthesiologica Scandinavica*, 50(3):368–373, 2006.

[20] H. Blum. *Models for the Perception of Speech and Visual Form*. Cambridge, MA: MIT Press, 1967.

[21] P. V. C. Hough. Machine analysis of bubble chamber pictures. International Conference on High Energy Accelerators and Instrumentation, CERN, 1959.

[22] R. M. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Transactions on Systems, Man and Cybernetics*, 3(11):610–621, 1973.

[23] M. Bevk and I. Kononenko. A statistical approach to texture description of medical images: A preliminary study. In *The 19th International Conference on Machine Learning ICML'02 Workshops*, Sydney, 2002.

[24] M. Tuceryan and A. K. Jain. Texture segmentation using Voronoi polygons. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 211–216, 1990.

[25] M. Tuceryan and A. K. Jain. Texture analysis. In C. H. Chen, L. F. Pau, and P. S. P. Wang, Eds., *The Handbook of Pattern Recognition and Computer Vision*, 2nd ed., pp. 207–248. River Edge, NJ: World Scientific, 1998.

[26] A. Rosenfeld and A. C. Kak. *Digital Picture Processing*, 2nd ed., vol. 2. New York: Academic Press, 1982.

[27] J. A. Rushing, H. S. Ranganath, T. H. Hinke, and S. J. Graves. Using association rules as texture features. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(8):845–858, 2001.

[28] K. I. Laws. *Textured Image Segmentation*. PhD thesis, Dept. Electrical Engineering, University of Southern California, 1980.

[29] B. S. Manjunath and W. Y. Ma. Texture features for browsing and retrieval of image data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(8):837–842, 1996.

[30] R. Agrawal, T. Imielinski, and A. N. Swami. Mining association rules between sets of items in large databases. In Peter Buneman and Sushil Jajodia, Eds., *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data*, pp. 207–216, New York: ACM Press: 1993.

[31] S. Brin, R. Motwani, and C. Silverstein. Beyond market baskets: Generalizing association rules to correlations. In Joan Peckham, Ed., *SIGMOD 1997, Proceedings ACM SIGMOD International Conference on Management of Data, May 13–15, 1997, Tucson, Arizona, USA*, pp. 265–276. New York: ACM Press, 1997.

[32] P. Tan and V. Kumar. Interestingness measures for association patterns: A perspective. Technical Report TR00–036, Department of Computer Science, University of Minnesota, 2000.

[33] P. Tan, V. Kumar, and J. Srivastava. Selecting the right interestingness measure for association patterns. In *Proceedings of the Eight ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2002.

[34] M. Bevk. *Derivation of texture features using association rules*. PhD thesis, Faculty of Computer and Information Science, University of Ljubljana, 2005. In Slovene.

[35] C. Bastos Rocha Ferreira and D. Leandro Borges. Automated mammogram classification using a multiresolution pattern recognition approach. *SIBGRAPI01*, 00:76, 2001.

[36] M. L. Comer and E. J. Delp. Segmentation of textured images using a multiresolution Gaussian autoregressive model. *Image Processing, IEEE Transactions on Image Processing*, 8(3):408–420, 3, 1999.

[37] T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution grayscale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):971–987, 2002.

[38] C. K. Chui. *An Introduction to Wavelets*. San Diego: Academic Press, 1992.

[39] K. Laws. *Textured image segmentation*. PhD thesis, University of Southern California, 1, 1980.

[40] S. E. Grigorescu, N. Petkov, and P. Kruizinga. Comparison of texture features based on Gabor filters. *IEEE Transactions on Image Processing*, 11(10):1160–1167, 2002.

[41] I. H. Witten and E. Frank. *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*. San Francisco: Morgan Kaufmann, 1999.

[42] X. Shen, M. Boutell, J. Luo, and C. Brown. Multi-label machine learning and its application to semantic scene classification. In *Proceedings of the 2004 International Symposium on Electronic Imaging (EI 2004)*, San Jose, California, 2004.

[43] A. McCallum. Multi-label text classification with a mixture model trained by EM. In *Proceedings of the AAAI'99 Workshop on Text Learning*, 1999.

AQ:1 do you mean minus, or should this be=?
AQ:2 correct?
AQ:3 correct? or public, as in last sentence of paragraph?
AQ:4 sentence ok? should image processor be lowercase? should "program" be "algorithm"?
AQ:5 wording ok?
AQ:6 change correct?
AQ:7 change correct?
AQ:8 check wording?
AQ:9 Change correct?